

HOUSE PRICE PREDICTION USING MACHINE LEARNING

V. SARALA ,G. PAVANI

Assistant Professor MCA, DEPT, Dantuluri Narayana Raju College, Bhimavaram, AndhraPradesh

Email id:- vedalasarala21@gmail.com

PG Student of MCA, Dantuluri Narayana Raju College, Bhimavaram, AndhraPradesh

Email id:- Pavanipretty6788@gmail.com

ABSTRACT

The sales of the houses are determined on various factors like the location, area, population and some of the information to predict the individual housing price. In addition to these housing prices, the prediction of the housing prices can greatly assist in the prediction of the future housing prices of the real estate. This study uses the machine learning algorithms and technology as a research methodology to develop a housing price prediction. Many algorithms are used here to effectively increase the percentage of the prediction which is considered as the best models in the price prediction. This project shows us that the machine learning algorithm based on accuracy, consistency out performs the other in the performance of the housing price prediction.

The project can be created using python(AI/ML),HTML,CSS,JSS.Python is used for writing the Machine Learning Algorithms .HTML,CSS and JS is used for designing the front end of the system.

At last, I can conclude y saying that House Price Prediction system will be very helpful in detecting the prices of the houses and keeping the record of the high and low of the prices. So it will help the user to know the real price of the property, it could not be used for any fraud means.

1. INTRODUCTION

Earlier, it's a very popular and common practice to price the property without the proper evaluation of the land, infrastructure etc. We need a proper prediction on the real estate and the houses in housing market we can see a mechanism that runs throughout the properties buying and selling buying a house will be a life time goal for most of the individual but, There are lot of people making mistake in India as most of the people are buying properties from the people they don't know by seeing the news all around them. In India, people buy properties which are too expensive but it's not worth it.

In the housing market 2016 the house sold in India was about 80 lakh but the real price according locality and size was about 60 lakh.

In earlier year, there was an economic collapse that give the clue to the impending disaster, this situation is currently happening and the prices of houses are getting higher compared to current economic situation of our country, the Indian government fails to produce the data about the houses so it was very difficult for peoples to buy the properties. Therefore, the people searched on internet for the evidence for house price.

Many methods have been used in the price prediction like advance regression in this I am trying to predict the real estate price for the future using the machine learning techniques with the help of the previous works. I have used the multiple regression and more algorithms with different tools to predict the house price. The purpose of this paper is to establish the proper data preprocessing practices in order to increase the accuracy of machine learning algorithms

2. LITERATURE SURVEY AND RELATED WORK

The real estate market is one of the most competitive in terms of pricing and same tends to be vary significantly based on lots of

factor, forecasting property price is an important modules in decision making for both the buyers and investors in supporting budget allocation, finding property finding stratagem and determining suitable policies hence it becomes one of the prime fields to apply the concepts of machine learning to optimize and predict the prices with high accuracy.

The literature review give the clear idea and it will serve as the support for the future projects. Most of the authors have concluded that artificial neural network have more influence in predicting the but in real world there are other algorithms which should have taken into the consideration. Investors decisions are based on the market trends to reap maximum returns. Developers are interested to know the future trends for their decision making, this helps to know about the pros and cons and also help to build the project. To accurately estimate property prices and future trends, large amount of data that influences land price is required for analysis, modelling and fore casting. The factors that affect the land price have to be studied and their impact on price has also to be modelled. It is inferred that establishing a simple Regression linear mathematical relationship for these time-series data is found not viable for prediction. Hence it became imperative to establish a non-linear model which can well fit the data characteristic to analyze and predict future trends. As the real estate is fast developing sector, the analysis and prediction of land prices using mathematical modelling and other techniques is an immediate.

3. EXISTING SYSTEM

Timothy C. Au [13] addressed about the absent level problems in Random Forests, Decision Trees, and Categorical Predictors. Using three real data sets, the authors have illustrated how the absent levels affect the performance of the predictors.

There are many existing approaches that can be used to determine the prices of the house, one of them is prediction analysis.

first approach looks for the time-series data. The time-series approach is to look for the relationship between current prices and prevailing prices. The existing system calculates the price of the house without knowing the information for the future and necessary prediction. This project House Price Prediction helps the people who wanted to buy the house so they can know the price range in the future.

House price prediction also helps the property dealer to know the worth of the property in future.

The main aim or focus of our project is to predict the accurate price of the real estate properties present in India for the next upcoming years through different Algorithms.

A.) Linear Regression

It is a supervised learning technique and responsible for predicting the value of a dependant variable (Y) based on the given independent variable (X).It is the relationship between the input(X) and output(Y) .

B.) Multiple correlation Analysis

It helps to take out the maximum degree of linear relationship that can be obtained between two or more independent variables and a single dependant variable

C.) Classification Trees

Classification Trees are used to predict the object into classes of a categorial dependant variable based on the one or more predictor variables.

Technology used

A.) Data Science :-

Data science is the first stage in which we take the dataset and will do the data cleaning on it. We will do the data cleaning to make sure that it provides the reliable predictions.

B.) Machine Learning:-

The cleaned data is fed into the machine learning model, and we do some of the algorithms like

C.) Front End (UI)

The front end is basically the structure or a build up for a website. In this to receive an information for predicting the price .

It takes the form data entered by the user and executes the function which employs the prediction model to calculate the predicted price for the house.

4. PROPOSED SYSTEM

Proposed work aims at predicting the availability of houses based on different features of the houses and also the facilities available nearby the location of the houses. Work also includes the price prediction of the houses based on the features of the house and facilities nearby its location

i)Decision Tree Classifier is used to predict the availability of houses as per the users' requirement constraints and it produces responses like yes or no respectively to tell whether a house is available or not.

(ii)Decision tree regression and Multiple Linear Regression methods are used to predict the prices of the houses. A real time dataset is prepared by analyzing the location named Tadepaligudem of West Godavari District in Andhra Pradesh of India. The dataset contains the following features of the houses such as Number of bedrooms, age of the house, transport facility, schools available in the nearby location and shopping facilities.

The proposed method helps to search houses in big cities based on the following attributes.

1. Number of bedrooms (1BHK, 2BHK and 3BHK).
2. Transport facility such as availability of bus facility, train facility and flight facility.
3. School facility such as availability of Government schools, matriculation and CBSE.
4. Shopping facility such as small markets, general stores, shopping malls
5. Prices of the houses from 10 lakhs to 30 lakhs.
6. Age of the house varying from one to five years. The proposed work is implemented using Scikit Learn, a machine learning tool

5. METHODOLOGIES

The main goal of the project is to find out the accurate predictions of the houses/ properties present in the

Bengaluru for the next upcoming years. Here are the step by step process involved

1. Requirement Gathering – We have to gather the information extract the main information from it.
2. Normalizing the data
3. Detecting Outliners in the data
4. Analysis and visualization using the data

The Extracted Details :

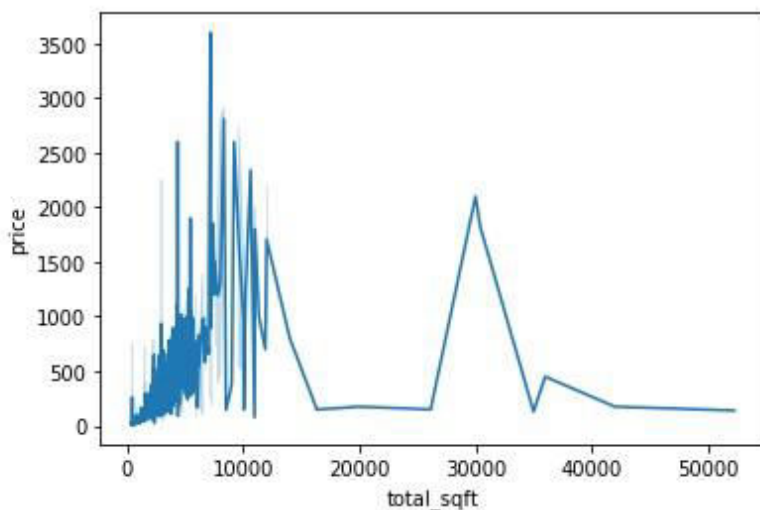
```
In [42]: data.head()
```

```
Out[42]:
```

	location	size	total_sqft	bath	price	bhk	price_per_sqft
0	Electronic City Phase II	2 BHK	1056.0	2.0	39.07	2	3699.810606
1	Chikka Tirupathi	4 Bedroom	2600.0	5.0	120.00	4	4615.384615
2	Uttarahalli	3 BHK	1440.0	2.0	62.00	3	4305.555556
3	Lingadheeranahalli	3 BHK	1521.0	3.0	95.00	3	6245.890861
4	Kothanur	2 BHK	1200.0	2.0	51.00	2	4250.000000

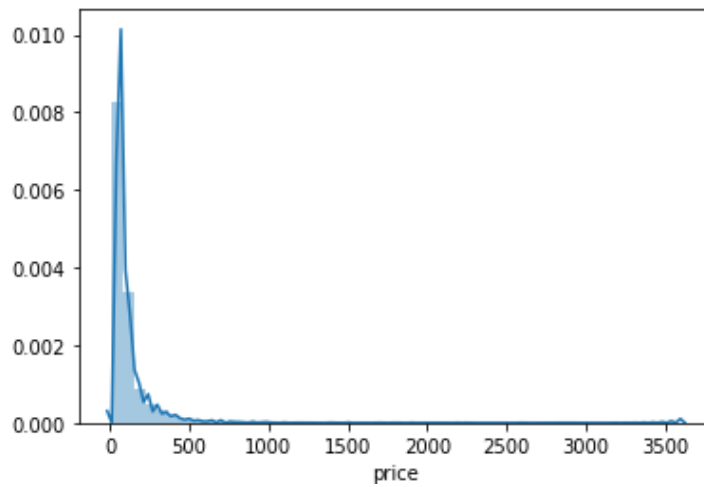
Project Graph-

```
In [81]: sns.lineplot(x="total_sqft",y="price",data= data)
plt.show()
```



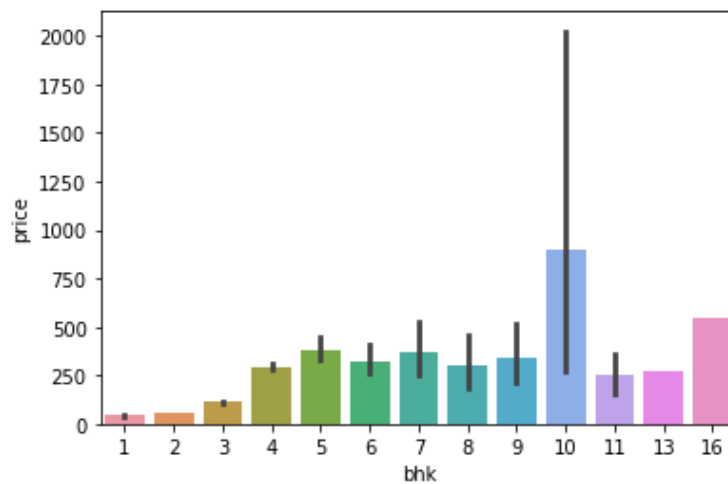
```
In [51]: sns.distplot(data['price'])
```

```
Out[51]: <matplotlib.axes._subplots.AxesSubplot at 0x1509399cd08>
```



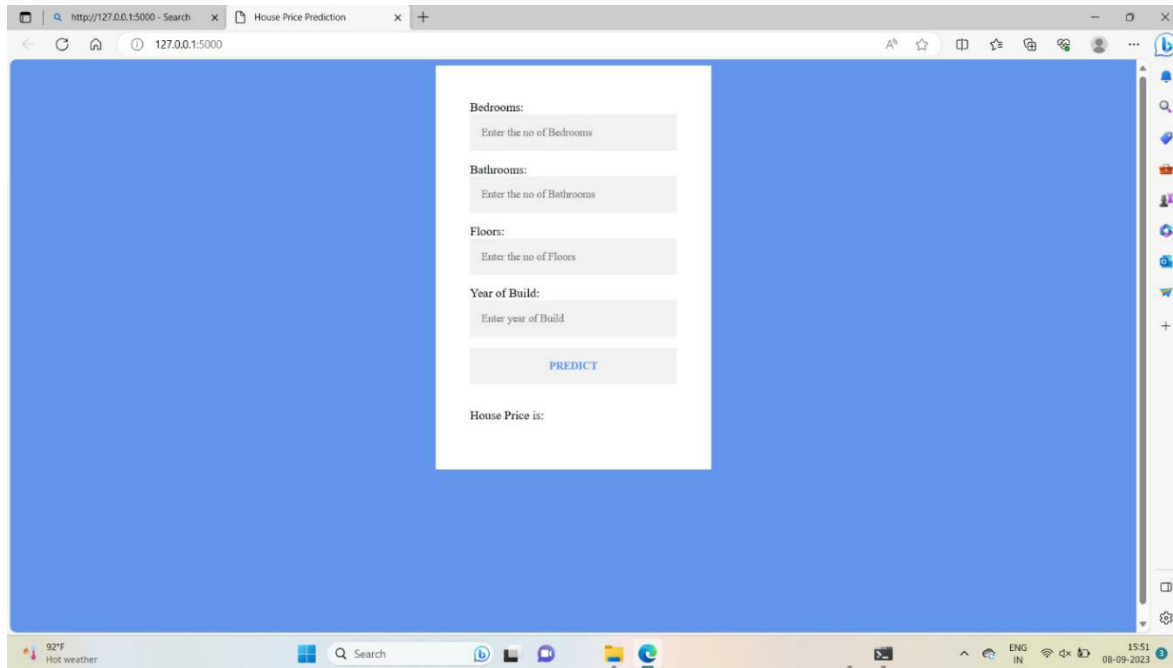
```
In [80]: sns.barplot(x="bhk", y="price", data=data)
```

```
Out[80]: <matplotlib.axes._subplots.AxesSubplot at 0x1de77c58748>
```

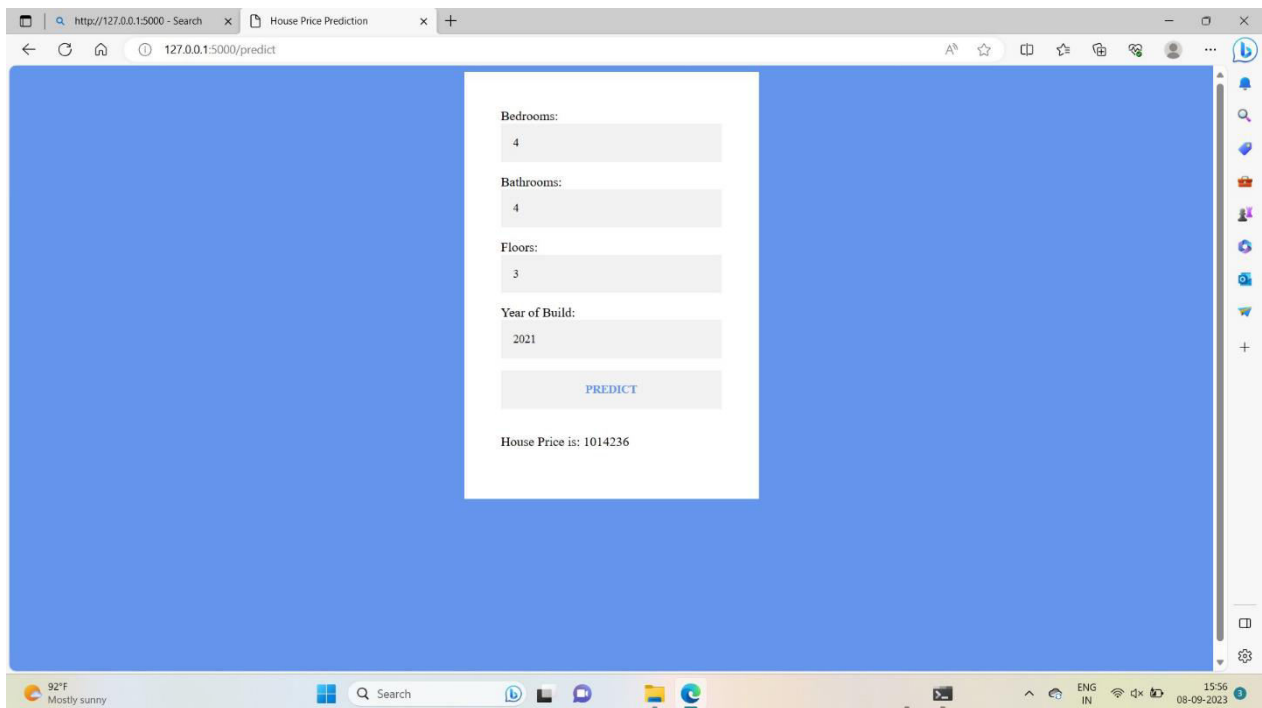


6. RESULTS AND DISCUSSION SCREEN SHOTS

HOME SCREEN :



OUTPUT SCREEN :



7. CONCLUSION AND FUTURE SCOPE

The main aim of this project is to determine the prediction of prices. In this paper, we have discovered many algorithms and application of machine learning techniques with the objective of buying the real estate properties and to predict the worth in the future of the owned real estate properties.

Price can be predicted through many factors like the surrounding, marketplaces and many related factors with the house. We have first cleaning and exploring of the input data. The predicted data can be stored in the database and app or website can be made for the people, as people can have the brief idea of the property.

We have performed ensembles of regression trees, k-nearest neighbors, multi-linear regression as we understand that parameterization of the can drive the significant result in the performance

8. REFERENCES

1. Ganjisaar, Y., Caruana, R. and Lopes, C. V. (2011). Bagging gradient-boosted trees
2. Chang, P.-C. and Liu, C.-H. (2008). A tsf type fuzzy rule based system for stock price prediction, Expert Systems with applications
3. Ganjisaar, Y., Caruana, R. and Lopes, C. V. (2011). Bagging gradient-boosted trees for high precision, low variance ranking models, Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval, ACM,
4. Gu, J., Zhu, M. and Jiang, L. (2011). Housing price forecasting based on genetic algorithm and support vector machine, Expert Systems with Applications Koskela, T., Lehtokangas, M., Saarinen, J. and Kaski, K. (1996). Time series prediction with multilayer perceptron, radial basis function and elman neural networks, Proceedings of the World
5. Aswin Sivam Kumar Research Paper (Real Estate Price Prediction Using Machine Learning, School of computing National College of Ireland)